

A formulák az ELTE ösztintézeti szemináriumi előadásomba

Varga Dániel

March 26, 2018

Rényi Intézet

- $f_{\Theta} : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Folytonos, majdnem mindenhol differenciálható függvények családja, amelynek egyes elemeit valós paraméterek Θ vektora jelöli ki.
- A függvény által elvárt viselkedést az $(x_1, y_1), \dots, (x_N, y_N)$ példákkal (tanítókorpusszal) adjuk meg. $x_i \in \mathbb{R}^n, y_i \in \mathbb{R}^m$.
- Továbbá az $L(\hat{y}, y)$ veszteségfüggvénnyel, amely megadja, hogy mennyire vagyunk elégedettek az $\hat{y}$ kimenettel, amikor y lenne az elvárt kimenet. (Konvenció szerint kisebb érték a jobb.)

A fenti jelölésekkel

$$\operatorname{argmin}_{\Theta} \sum_{i=1}^N L(f_{\Theta}(x_i), y_i)$$

adja meg az optimális paraméterbeállítást. Ha egy optimalizálási folyamat végén rögzítünk egy Θ paramétervektort, amely a fenti értéket kicsivé teszi, akkor azt mondjuk, hogy a hálózatot betanítottuk a tanítókorpuszon.

$$W_1(\mu, \nu) = \inf_{\gamma \in \Gamma(\mu, \nu)} \mathbb{E}_{(x,y) \sim \gamma} d(x, y),$$

ahol $\Gamma(\mu, \nu)$ a csatolások halmaza. A csatolás olyan mérték $M \times M$ felett, amelynek marginális eloszlásai az első és második koordinátán μ illetve ν .

$$W_1(\mu, \nu) = \sup_{f: M \rightarrow \mathbb{R}, \text{Lip}(f) \leq 1} (\mathbb{E}_{y \sim \nu} f(y) - \mathbb{E}_{x \sim \mu} f(x)),$$

ahol $\text{Lip}(f)$ az f függvény Lipschitz konstansa.

Ez egy klasszikus lineáris programozási dualitás, és ahogy ez ilyeneknél szokásos, az egyik irányú egyenlőtlenség könnyű, a másik nehéz.

Nem fogjuk bebizonyítani, de legalább nézzük meg a könnyű irányát.

Az alvállalkozó. (Caffarelli via Villani) A μ 1 tonnás homokkupacot kellene áthordanunk a ν pozícióba. Egy tonna homok x -ből y -ba való áthordásának költsége $d(x, y)$, a költség egyenesen arányos a tömeggel. Megjelenik egy gyanús vállalkozó, és vállalja, hogy áthordja a homokkupacunkat.

Rajzol nekünk egy f potenciálfüggvényt (domborzati térképet), azzal a tulajdonsággal, hogy két pont közti szintkülönbség sosem nagyobb, mint a távolságuk. Az elszámolás módjának az alábbi javasolja: ha m tonna homokot x -ből y -ba kellett szállítania, azért $(f(y) - f(x))m$ költséget számláz.

Az világos, hogy $\mathbb{E}_{y \sim \nu} f(y) - \mathbb{E}_{x \sim \mu} f(x)$ összeget fog kiszámlázni, és ez teljesen független attól, hogy ő maga hogyan oldotta meg a szállítást. Tehát feltehetjük, hogy a lehető leghatékonyabban járt el, és ez esetben az is világos, hogy nem járhatunk rosszul az ajánlatával, az f Lipschitz-tulajdonsága miatt.

A dualitástétel nehéz iránya azt mondja ki, hogy jól sem fogunk járni.